

Attachment A

Abstract:

5

A method of determining cluster attractors for a plurality of documents comprising at least one term. The method comprises calculating, in respect of each term, a probability distribution indicative of the frequency of occurrence of the, or each, other term that co-occurs with said term in at least one of said documents. Then, the entropy of the respective probability distribution is calculated. Finally, at least one of said probability distributions is selected as a cluster attractor depending on the respective entropy value. The method facilitates very small clusters to be formed enabling more focused retrieval during a document search.

15